

Lecture Notes in Computer Science

719

**Dmitry Chetverikov
Walter G. Kropatsch (Eds.)**

Computer Analysis of Images and Patterns

**5th International Conference, CAIP '93
Budapest, Hungary, September 1993
Proceedings**



Springer-Verlag

Dmitry Chetverikov Walter G. Kropatsch (Eds.)

Computer Analysis of Images and Patterns

5th International Conference, CAIP '93
Budapest, Hungary, September 13-15, 1993
Proceedings

Springer-Verlag

Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Series Editors

Gerhard Goos
Universität Karlsruhe
Postfach 69 80
Vincenz-Priessnitz-Straße 1
D-76131 Karlsruhe, Germany

Juris Hartmanis
Cornell University
Department of Computer Science
4130 Upson Hall
Ithaca, NY 14853, USA

Volume Editors

Dmitry Chetverikov
Computer and Automation Research Institute, Hungarian Academy of Sciences
POB 63, H-1518 Budapest, Hungary

Walter G. Kropatsch
Institut für Automation, Technische Universität Wien
Treitlstraße 3, A-1040 Vienna, Austria

CR Subject Classification (1991): I.4-5

ISBN 3-540-57233-3 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-57233-3 Springer-Verlag New York Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1993
Printed in Germany

Typesetting: Camera-ready by author
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
45/3140-543210 - Printed on acid-free paper

Planning the Next View Using the Max-Min Principle <i>J. Maver, A. Leonardis, F. Solina</i>	543
--	-----

Reconstruction from Projections

Cross-Correlation with Reconstruction: A New Approach to Pattern Matching <i>V. Shapiro</i>	548
The Iterated Normalized Backprojection Method of Image Reconstruction <i>W. L. Nowinski</i>	556
Cone Beam Reconstruction and Fourier Transform of Distributions <i>O. Trofimov</i>	564

Character Recognition and Document Processing

Character Recognition by Affine Moment Invariants <i>J. Flusser, T. Suk</i>	572
Ultra Fast Pattern Classification by Fuzzy Logic <i>R. Fageth, W. G. Allen, U. Jäger</i>	578
An Object-Oriented Pen-Based Recognizer for Handprinted Characters <i>B. Klauer, K. Waldschmidt</i>	586
On the Method of Critical Points in Character Recognition <i>E. V. Shchepin, G. M. Nepomnyashchii</i>	594
Printed Text Segmentation Using Distance Transform <i>S. Ablameyko, O. Okun</i>	599
Recognition of Handwritten Layout Drawings <i>S. Takamura, M. Takagi</i>	604
Automatic Recognition of Scanned Technical Drawings <i>F.-L. Krause, H. Jansen, N. Luth</i>	609

Biomedical Applications

Digital Image Processing in Radiology. An Experience of Development of Technology for Computer-Aided Diagnosis <i>T. P. Belikova (Invited)</i>	614
---	-----

Planning the Next View Using the Max-Min Principle

Jasna Maver, Aleš Leonardis, and Franc Solina
Computer Vision Laboratory, Dept. of Elec. Eng. and Comp. Science,
University of Ljubljana, 61001 Ljubljana, Tržaška 25, Slovenia

Abstract. We present in this paper an approach to a typical task of active sensing, namely how to successively position the sensor in order to accomplish a given task. The approach is general in the sense that it does not require *a priori* information about the constituents of the scene, and *data driven* in the sense that it is the data that provides the information for driving the planning of the next views. We applied our approach to the task of determining the 3-D coordinates of the points of object silhouettes in the image under orthographic projection.

1 Introduction

It has long been realized that information obtained from a single viewpoint might not be sufficient for successfully accomplishing a task and that additional views are necessary [1,2,4].

In this paper we present an active sensor system which tends to accomplish a given task by acquiring the minimal number of images. The principle which guides the next-view planning is the *Max-Min principle*. The viewing direction from which the first image is acquired is arbitrary or predefined. To select the new viewing direction for image acquisition the sensor system must estimate how much of the yet unknown data necessary to accomplish the task can be acquired from each possible viewing direction. This estimate is based on the information acquired in previous images. When there is insufficient information the worst-case situation is assumed. Thus for each viewing direction the minimal number of data that can be obtained is anticipated, i.e., the utilization of the minimum principle. To select the next best viewing direction we propose two different strategies:

1. Select from all possible viewing directions the one which gives under the worst-case assumptions for the unknown data the maximal amount of information.
2. First compute the necessary viewing directions from which all the necessary data to accomplish the task can be acquired, and then select among these viewing directions the one which maximizes the amount of new information.

While the first strategy can be compared to a greedy algorithm which tries to obtain at each step the maximal information regardless of the final number of views, the second strategy with its look-a-head capability takes into account a more global view.

2 Geometry of the viewer

Let us define a scene in a rectangular coordinate system (X, Y, Z) . The viewer's

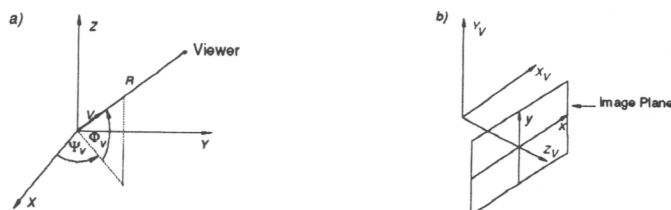


Fig. 1. (a) Viewing direction, (b) Camera image plane.

position is on a half sphere with radius R (Fig. 1(a)). The unit vector V pointing toward the viewer is given by $V^T = [\cos \Psi_V \cos \Phi_V, \sin \Psi_V \cos \Phi_V, \sin \Phi_V]$, where Ψ_V denotes the right ascension and Φ_V the elevation. The viewer's coordinate system is located in the center of the scene coordinate system and defined by the three axes (X_V, Y_V, Z_V) .

The coordinates of a point $P = (X_P, Y_P, Z_P)$ in the viewer's coordinate system are given by

$$\begin{bmatrix} X_{VP} \\ Y_{VP} \\ Z_{VP} \end{bmatrix} = \begin{bmatrix} -\sin \Psi_V & \cos \Psi_V & 0 \\ -\cos \Psi_V \sin \Phi_V & -\sin \Psi_V \sin \Phi_V & \cos \Phi_V \\ \cos \Psi_V \cos \Phi_V & \sin \Psi_V \cos \Phi_V & \sin \Phi_V \end{bmatrix} \begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix}. \quad (1)$$

The camera image plane is parallel to the viewer's X_V - Y_V plane where the image axes (x, y) are aligned with axes (X_V, Y_V) (Fig. 1(b)). We assume that the imaging is performed under the orthographic projection:

$$x = X_V, \quad y = Y_V. \quad (2)$$

3 Task definition

A set of objects lies on the horizontal base surface of the scene. Without loss of generality we can assume that the first image is acquired from the viewing direction $V_1 = (0, 0, 1)$, i.e., $\Psi_{V_1} = -\frac{\pi}{2}, \Phi_{V_1} = \frac{\pi}{2}$.¹ In the image obtained from the first viewing direction the vision system extracts the silhouettes of the objects². The task that we study in this paper can now be formulated as: *Select the set of interesting points on the detected silhouettes and determine their 3-D coordinates.*

¹ The scene coordinate system can always be redefined in such a way that the first view is the one we have chosen.

² We assume that our vision system can distinguish the background points in the image from the points that are projections of the objects.

4 3-D coordinates of the points on the detected silhouettes

To design an efficient algorithm we first analyze the geometrical information contained in the points on the detected silhouettes acquired by the first two views.

First view: Object silhouette in the image of the first view is represented by a contour Q . Let $p = (x_p, y_p)$, $p \in Q$ be a projection of the silhouette point $P = (X_P, Y_P, Z_P)$. Eqs. (1) and (2) yield: $X_P = X_{V_P} = x_p$, $Y_P = Y_{V_P} = y_p$. Z_P coordinate remains unknown after the first view.

Second view: In order to determine the exact position for the second view we study the general case. x_p coordinate is not a function of Z_P therefore, after the first view, the x coordinate of any point P can be predicted for the known Ψ_{V_2} . For the viewing direction (Ψ_{V_2}, Φ_{V_2}) all the points in the scene on the plane $Y = \tan(\Psi_{V_2})X + n$ have the same X_V coordinate, thus the same x coordinate in the image, if they are not occluded:

$$X_V = x = -\sin(\Psi_V)X + \cos(\Psi_V)Y = n \cos(\Psi_V) . \quad (3)$$

Z_P coordinate can be computed from the $y_p = Y_{V_P}$ coordinate. Let us assume that there exists a line $y_{\Psi_{V_2}, n} = \tan(\Psi_{V_2})x + n$ in the image of the first view which passes through only one object point p . In the new image acquired from the viewing direction (Ψ_{V_2}, Φ_{V_2}) the image of the point P lies somewhere on the line $x = n \cos(\Psi_{V_2})$ (eq. 3). P has the largest Z coordinate among the points at location (X_P, Y_P) and therefore its projection has the largest y coordinate of object points on the line $x = n \cos(\Psi_{V_2})$ [3].

If the projection of the object silhouette Q is a polygon then a unique solution is possible only for the vertices v_i of the convex hull of Q . The vertices v_i determine the set of interesting points.

For each vertex v_i we compute the *ascension angle* $\Psi(v_i)$ and the *elevation angle* $\Phi(v_i)$ [3] from which the next ascension and elevation direction are selected.

Information acquired after the first two views: Let us denote the largest y coordinate of the objects points on the line $x = n \cos(\Psi_{V_2})$ in the image of the second view as y_{max_2} . A point on the line $y_{\Psi_{V_2}, n}$ in the image of the first view has the largest Z coordinate when $(n \cos(\Psi_{V_2}), y_{max_2})$ is its projection. We name that coordinate the *worst-case* Z coordinate and denote it as Z_{w-c} . For each object point we know its (X_P, Y_P) coordinates and know that $Z_P \leq Z_{w-c}(X_P, Y_P)$.

The third and all the following viewing directions: Knowing the Z_{w-c} coordinate for each pair (X, Y) of the scene we can use this information to compute new ascension and elevation angles for vertices v_i . Let the two points p_1 and p_2 lie on the line $y_{\Psi_{P_1, P_2}, n}$ in the image of the first view (Fig. 2(a)). p_1 and p_2 are the projections of the surface points P_1 and P_2 , respectively. Let the point P_1 be in front of the point P_2 relative to the viewer. The scene at the location (X_{P_1}, Y_{P_1}) can be occupied for $0 \leq Z \leq Z_{w-c}(X_{P_1}, Y_{P_1})$ and the scene at the location (X_{P_2}, Y_{P_2}) for $0 \leq Z \leq Z_{w-c}(X_{P_2}, Y_{P_2})$. The possible y coordinates of P_1 and P_2 are determined by the intervals (Fig. 2(b)):

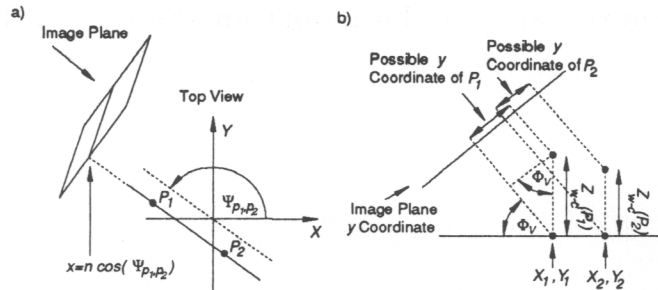


Fig. 2. Computation of the minimal viewing elevation.

$y((X_{P_1}, Y_{P_1}, 0), \Psi_{p_1, p_2}, \Phi_V) \leq y_{p_1} \leq y((X_{P_1}, Y_{P_1}, Z_{w-c}(X_{P_1}, Y_{P_1})), \Psi_{p_1, p_2}, \Phi_V)$,
 $y((X_{P_2}, Y_{P_2}, 0), \Psi_{p_1, p_2}, \Phi_V) \leq y_{p_2} \leq y((X_{P_2}, Y_{P_2}, Z_{w-c}(X_{P_2}, Y_{P_2})), \Psi_{p_1, p_2}, \Phi_V)$, respectively. The two intervals do not overlap if

$$\Phi_V > \arctan\left(\frac{Z_{w-c}(X_{P_1}, Y_{P_1})}{\sqrt{(Y_{P_1} - Y_{P_2})^2 + (X_{P_1} - X_{P_2})^2}}\right) = \Phi_{\min}(p_1, p_2) \quad (4)$$

For viewing direction $(\Psi_{p_1, p_2}, \Phi_V > \Phi_{\min}(p_1, p_2))$ P_1 does not occlude P_2 .

5 Determining the next viewing direction

For the next viewing direction we must determine the ascension and the elevation direction. The selection of ascension direction has a dominant role.

First we build a histogram $\mathcal{H}(\psi)$ which shows how many vertices can be determined for any ascension direction $\psi \in [0, 2\pi)$.

One possible solution is to select as the next ascension direction the one from which the largest number of vertices can be determined, i.e., ψ for which $\mathcal{H}(\psi)$ has a global maximum.

Another possible solution to the problem is first to analyse the histogram $\mathcal{H}(\psi)$ and find the necessary number of viewing directions from which we can compute the Z coordinate of all the vertices and then select among them the one from which the largest number of vertices can be determined. Due to the lack of space the reader is referred to [3].

6 Experimental Results

A scene consists of two polyhedra and two pyramids of different heights and sizes. The first viewing direction is $\Psi_{V_1} = -\pi/2, \Phi_{V_1} = \pi/2$ (Fig. 3(a)). We locate 16 vertices. For each vertex we compute the ascension angle and build the histogram (Fig. 3(b)). For Ψ_{V_2} we choose the value from one of the histogram maxima: $\Psi_{V_2} = 270, 5^\circ$ and $\Phi_{V_2} = 76.5^\circ$. From this view we can determine the Z coordinates of six vertices marked by circles in Fig. 3(c). Fig. 4(a) shows the image taken from the selected view. Fig. 4(b) depicts the y_{max_2} coordinates from which we can get the Z_{w-c} coordinates.

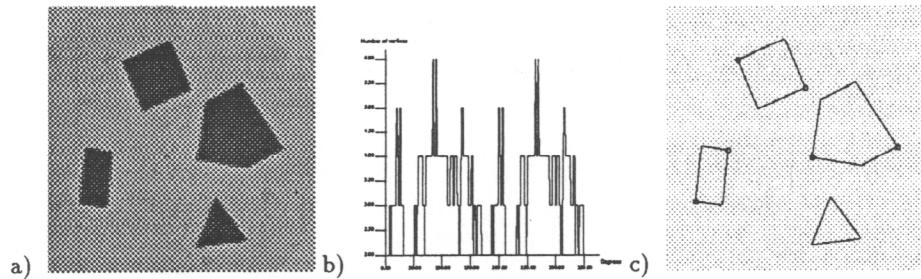


Fig. 3. (a) Image of the scene taken from the first view, (b) Histogram showing the number of vertices for which the Z coordinate can be determined with respect to the ascension angle, (c) Six selected vertices.

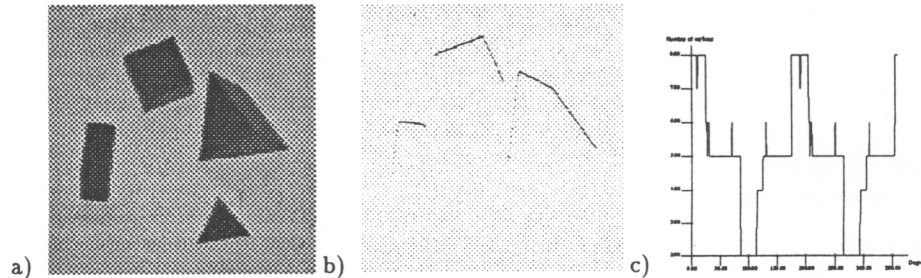


Fig. 4. (a) The image of the scene taken from the second viewing direction, (b) y_{max2} coordinates, (c) Histogram of remaining 10 vertices.

We compute new ascension angles for the remaining 10 vertices. Fig. 4(c) depicts the new histogram. We find that two additional images must be acquired to accomplish the task: one ascension direction must be selected from the interval $(175^\circ, 189^\circ)$ or $(355^\circ, 9^\circ)$, another from the interval $(27^\circ, 85^\circ)$ or $(207^\circ, 265^\circ)$. We choose as a third ascension direction $\Psi_{V_3} = 2^\circ$. For the vertices v_i which are selected by this ascension the minimal elevation $\Phi_{min}(v_i)$ has to be computed. Elevation direction must be $\Phi_{V_3} > \max_i(\Phi_{min}(v_i))$. We compute: $\Phi_{V_3} > 67.3^\circ$.

References

- [1] J.Aloimonos and A.Bandopadhyay, "Active Vision," in *IEEE 1st Int. Conf. on Computer Vision*, pp. 35-54, June 1987.
- [2] R.Bajcsy, "Active Perception", *Proceedings of the IEEE*, August 1988.
- [3] J.Maver, "Planning the Next View Using the Max-Min Principle," University of Ljubljana, Faculty of EE and CS, Tech. report LRV-93-4, 1993.
- [4] J.Maver and R.Bajcsy, "Occlusions as a Guide for Planning the Next View," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, May 1993.